

ANÁLISIS DE TRES MEDIDAS DE CARGA MENTAL: JUICIOS SUBJETIVOS, TAMAÑO DE LA PUPILA Y TASA DE PARPADEOS

*Elisa Pérez, Miguel Ángel Recarte y
Ángela Conchillo*

*Facultad de Psicología
Universidad Complutense de Madrid*

INTRODUCCIÓN

El concepto de carga mental tiene definiciones poco operativas derivadas de una multiplicidad de conceptos psicológicos más primitivos: existen definiciones basadas en las etapas de procesamiento de información (De Waard, 1996, 2002), en la distinción entre procesos automáticos y controlados (Mulder, 1980), en el concepto de esfuerzo (Vicente, Thornton y Moray, 1987), o en el de capacidad limitada (O'Donnell y Eggemeier, 1986). Aún así, parece existir un acuerdo en que la carga mental no viene derivada únicamente de propiedades de las tareas, ni de características del operador, sino de interacciones entre ambas. El interés por su medición se deriva de que se ha hecho de ella un constructo explicativo del rendimiento en situaciones y tareas en que el ser humano se caracteriza como un operador, y en las que una parte importante de los fallos en la ejecución conlleva situaciones de peligro; la asunción subyacente es que tales fallos tienen que ver con sobrecarga del operador. Los campos donde más activamente se han estudiado las consecuencias de la carga mental inadecuada quizá sean el de en-

trenamiento de pilotos para el tráfico aéreo (Wickens, 1992) y el del tráfico de carretera (De Waard, 1996).

El presente trabajo ha surgido precisamente del ámbito de la investigación de la carga mental aplicada a la conducción de automóviles. Recarte y Nunes (2000) investigan los efectos, sobre los patrones de movimientos oculares, de la carga mental producida por la realización de tareas mentales mientras se conduce. Recarte y Nunes (2003) replican el estudio anterior y sus resultados con otras tareas mentales y extienden sus efectos a las capacidades para detectar y discriminar estímulos, y Recarte y Nunes (2001) estudian los efectos de la carga mental y del tiempo conduciendo sobre la tasa de parpadeo. Son estos trabajos, realizados todos ellos en conducción real ordinaria con un vehículo instrumentado y equipado con un sistema de registro de las miradas, los que han guiado la selección de variables independientes y tareas realizadas, la selección de las medidas de la carga mental y el diseño experimental, en un intento de consolidar en el laboratorio algunos de los resultados obtenidos en conducción real, así como la de verificar algunas hipótesis referidas al concepto de carga mental.

En correspondencia con el ambiguo estatus de las definiciones, las medidas de carga mental son muy variadas. O'Donnell y Eggemeier (1986) las agrupan en tres clases: medidas subjetivas o de autoinforme, medidas de ejecución y medidas fisiológicas. De Waard (1996), siguiendo esta clasificación, discute las características de las medidas. También pueden verse comentarios críticos en Sanders (1998) y en Wickens (1992). En el presente trabajo nos proponemos analizar tres de esas medidas, tomadas a partir de la ejecución de unas mismas tareas, tratando de estudiar en qué aspectos son equivalentes y en cuáles muestran divergencias indicativas de que miden diferentes aspectos del constructo de carga mental. Concretamente, vamos a analizar conjuntamente: (a) una medida de autoinfor-

me: NASA Task Load Index (NASA TLX), y dos medidas fisiológicas: (b) la dilatación de la pupila y (c) la tasa de parpadeo. Además, consideraremos la relación de estas medidas con la ejecución en una tarea de búsqueda visual continua dado que la carga mental ha sido un constructo explicativo del deterioro en ejecución de tareas.

Algunos autores piensan que la medida más precisa de la carga mental implicada en la realización de una tarea es preguntárselo al operador (De Waard, 1996) y, de hecho, en algunos trabajos de validación de medidas fisiológicas de carga mental se utiliza como criterio una medida subjetiva (Iqbal, Zheng y Bailey, 2004). De Waard (1996) habla de la posibilidad de que los sujetos tengan acceso introspectivo directo a alguna clase de representación de la carga mental, pero analiza las críticas que ha recibido esta medida, como la dificultad de diagnosticar introspectivamente las diferentes demandas que supone la realización de una o varias tareas. Nuestro diseño, con tareas simples y dobles, va a permitir estudiar este aspecto. En nuestro caso utilizaremos la escala NASA TLX (Hart y Staveland, 1988). Según Rubio, Díaz, Martín y Puente (2004), la sensibilidad del instrumento ha sido demostrada utilizando una gran variedad de tareas, resaltando sus componentes multidimensionales, que proveen información diagnóstica sobre fuentes diferentes de carga mental, así como su escasa o nula intrusión en la realización de las tareas y su validez para predecir la ejecución de los sujetos en diferentes tareas. NASA TLX es una escala multidimensional con 6 subescalas: Exigencia Mental, Exigencia Física, Exigencia Temporal, Esfuerzo, Rendimiento y Frustración. Se administra en dos etapas: en la primera, los sujetos evalúan, mediante el procedimiento de comparaciones por pares, cuánto contribuye cada una de las 6 subescalas a la definición de carga de trabajo, sin conocer las tareas que van a ser evaluadas. Las puntuaciones,

así obtenidas, se usarán como coeficientes de ponderación. En la segunda fase, el sujeto juzga, en una escala de 0 a 20, cuál es la contribución de cada subescala a la carga de trabajo de la tarea evaluada. Estos juicios son ponderados por los coeficientes de la primera fase para obtener una puntuación global.

La dilatación de la pupila ha sido considerada como un indicador fiable y sensible de carga mental (Beatty, 1982; Hoecks y Levelt, 1993). Un problema con su utilización es que su dilatación no es un indicador específico de carga mental, pues, junto a su función básica de regular la cantidad de luz que incide sobre la retina, también se dilata ante la presencia de estímulos con connotaciones emocionales o en situaciones en las que se produce un reflejo de orientación (Näätänen, 1992). No obstante, a pesar de las reticencias de algunos investigadores a favorecer el uso de la dilatación de la pupila como una medida de carga mental, existen claros resultados, en el contexto de la conducción de automóviles, que lo apoyan. Recarte y Nunes (2000) muestran la sensibilidad de la pupila para reflejar la realización de tareas mentales en periodos de 30 segundos, durante la conducción real en más de 100 Km. de carreteras y autovías. Estos resultados nos han llevado a estudiar la dilatación de la pupila con las mismas tareas en las que mostró diferencias en conducción real, en la hipótesis de que la pupila responde a la carga puramente cognitiva de la tarea.

De Waard (1996) incluye la tasa de parpadeo entre las medidas de carga mental. Otros autores estudian el parpadeo como indicador de fatiga (Stern, Boyer y Schroeder, 1994). Precisamente, este aspecto ha servido para los numerosos intentos de implementar en los vehículos dispositivos capaces de detectar los estados de fatiga y somnolencia de los conductores. Fogarty y Stern (1989) sostienen que la tasa de parpadeo guarda una relación inversa con las demandas visuales de la tarea, lo que puede explicarse como un

intento de minimizar la probabilidad de pasar por alto información importante. Recarte y Nunes (2001), con datos de conducción real, encuentran que, si bien la tasa de parpadeo aumenta con el tiempo conduciendo, es mayor el incremento atribuible a la realización de tareas mentales mientras se conduce. Esto plantea un problema de interpretación, pues dos situaciones, a primera vista opuestas en cuanto a la activación, producen el mismo resultado.

Los autores, a partir de las sugerencias de Stern et al. (1994), hipotetizan que la conducción, siendo una tarea con un fuerte componente visual, inhibe el parpadeo para no perder información visual del entorno. Tal inhibición del parpadeo supone un gasto de recursos, pues se trata en definitiva de una operación de control inhibitorio. La conducción prolongada (cansancio) hace disminuir los recursos disponibles y en parte se detraen del control del parpadeo, que de esta forma se des-inhibe regresando a su tasa natural, más frecuente. Por otra parte, la realización de tareas mentales también consume recursos, que en parte son detraídos de los utilizados para la inhibición del parpadeo; de este modo, el resultado final es el mismo que con el cansancio: des-inhibición e incremento del parpadeo para volver a su tasa natural. En nuestro experimento se diseñan las condiciones para probar en el laboratorio la doble hipótesis: (a) con tareas que requieren búsqueda visual continua se espera que disminuya la tasa de parpadeo; (b) con tareas cognitivas sin componente visual se espera un incremento del parpadeo proporcional a la carga que implica la tarea.

Recarte y Nunes (2003) encuentran que la mera comprensión verbal requiere poco esfuerzo, a juzgar por la dilatación de la pupila y por los juicios de esfuerzo asociados con su realización, mientras los sujetos conducen en tráfico real. Tampoco produce efectos en la realización de una tarea secundaria de detección de estímulos luminosos dentro del vehículo. Sin embargo, durante una tarea de

producción verbal se encuentra una clara dilatación de la pupila y un juicio subjetivo de esfuerzo alto. Coherentemente con esta carga mental incrementada se produce un decremento en el número de estímulos luminosos detectados. No obstante, Kubose et al. (2006) no encuentran diferencias entre comprensión y producción verbal tomando como medidas varios parámetros de conducción simulada. Por ello, creemos de interés replicar los anteriores resultados en condiciones más controladas y con más medidas que, en paralelo, confirmen la diferencia entre ambas tareas en cuanto a carga mental, por lo que se van a utilizar las mismas tareas de adquisición verbal y producción verbal. Acompañando a estas dos tareas cognitivas se ha incluido otra de cálculo o aritmética mental continua, también utilizada en los experimentos previos de conducción real, con carga mental similar a la producción verbal.

Finalmente, se incluye otra tarea que es clave para el estudio de las medidas: una tarea de búsqueda visual continua. Según nuestra hipótesis, esta tarea, por su diseño visualmente demandante, implica fuerte carga visual (búsqueda continua de una letra-estímulo entre 16 letras rápidamente cambiantes), pero poca carga cognitiva (mera identificación de la letra y traslación a respuesta pulsando una tecla). En consecuencia, se predice poca dilatación de la pupila y disminución del parpadeo por tratarse de una tarea de búsqueda visual, así como que la NASA TLX le asigne un valor intermedio de carga de trabajo, pues contiene una subescala que evalúa presión de tiempo para realizar la tarea. Esta tarea, en combinación con las anteriores, sirve también para introducir en el diseño el paradigma de doble tarea, un componente clásico en la investigación de las limitaciones de capacidad atencional.

En resumen, el presente trabajo trata de analizar tres medidas de carga mental: una escala subjetiva de carga de trabajo (NASA TLX), la

dilatación de la pupila y la tasa de parpadeo, tomadas durante la realización de tres tareas mentales y una tarea de búsqueda visual, en condiciones de tarea única o de doble tarea. Partimos de la idea de que el concepto de carga mental es un constructo para explicar la ejecución en una tarea. Los criterios principales de validez serán: la capacidad para discriminar entre tarea simple y doble, la capacidad para discriminar entre tareas cognitiva y tarea visual y, por último, la capacidad de las medidas para discriminar entre las tres tareas cognitivas. Por otra parte, se analizarán también las correlaciones entre las tres medidas. Finalmente, estudiaremos si el deterioro en la búsqueda visual puede ser predicho por las tres medidas de carga mental.

MÉTODO

Participantes

Los participantes del experimento fueron 48 personas de ambos sexos, con edades comprendidas entre los 18 y los 30 años (media: 20,52 años), cuarenta de los cuales, pertenecían al primer curso de la Facultad de Psicología de la Universidad Complutense de Madrid y su colaboración formó parte de su formación en las prácticas de la asignatura de Psicología de la Atención. El resto de los participantes fueron alumnos de otros cursos y materias de la misma facultad.

Diseño

Las variables independientes del experimento son: Tarea Cognitiva, con cuatro condiciones (Control, Escuchar, Producir palabras y Calcular) y Demanda Visual, con dos condiciones (No Detección y Detección), lo que da lugar a ocho condiciones experimentales en un diseño 4 (Tarea Cognitiva) x 2 (Demanda Visual) de medidas repetidas.

En la condición Control el sujeto no hacía la tarea, en la de Escuchar el sujeto oía una descripción grabada con instrucciones de que atendiera, pues posteriormente se le examinaría sobre lo escuchado; en la condición de Producir palabras los sujetos debían realizar una tarea de palabras encadenadas, a partir de una palabra dada por el experimentador debían generar palabras de forma que cada una comenzase por la última sílaba de la anterior; en la condición de Calcular o Aritmética Mental los participantes debían calcular en voz alta nuevas cifras, a partir de una cifra dada durante las instrucciones, siguiendo la norma de restar tres a la anterior y sumar siete verbalizando todos los resultados parciales.

En la condición de No Detección el sujeto debía mirar permanentemente a la pantalla en la que aparecían 16 letras cambiantes. En la condición Detección el sujeto debía realizar una tarea búsqueda visual en la pantalla consistente en detectar una letra (R, T, M o S, según la condición) entre 16 letras rápidamente cambiantes y pulsar la tecla “enter” de un teclado.

El término tarea doble se utiliza para la condición de Detección o Búsqueda Visual más hacer una de las tareas cognitivas.

Como variables dependientes, se utilizan: los Juicios en la escala NASA TLX, el Tamaño de la Pupila, el Número de Parpadeos por minuto (tasa de parpadeo) y el Rendimiento (proporción de aciertos) de los sujetos en la tarea de detección.

Materiales y procedimiento

El experimento se realizó en una cabina Faraday. El sujeto se sentó a 230 cm. frente a una pantalla que sirvió para la proyección de letras y que se utilizó como fondo visual a lo largo de todo el experimento. La pantalla subtendía 35,23° de ángulo visual horizontal y 26,90° de ángulo visual vertical. A una distancia de 94 cm. del ojo del sujeto, sin interferir con su línea de mirada a la pantalla, se colocó la cámara del sistema de movimientos oculares. Para evitar los movimientos

laterales de la cabeza los sujetos se apoyaron sobre una almohadilla de forma cóncava. Dos altavoces fueron utilizados para presentar las instrucciones del experimento. En otra zona de la cabina se situó el sistema de registro de movimientos oculares (Eye Tracking System, Model 5000 Control Unit, A.S.L), conectado a un segundo ordenador con el software específico del experimento, que presentaba, en modo audio, tanto las instrucciones generales como las de cada condición experimental, estas últimas balanceadas a través de los sujetos. Para contabilizar los parpadeos una cámara de vídeo grabó la imagen del ojo del participante que devolvía la cámara del sistema de movimientos oculares. Por último, se utilizó un luxómetro para medir la cantidad de luz en las sesiones. El nivel de luz a nivel del ojo fue de 80 lux y a nivel de la pantalla fue de 110 lux. El hecho de que los sujetos mirasen a esta pantalla durante las distintas fases aseguró una iluminación constante, esencial para el registro de la pupila.

El procedimiento para la realización del experimento fue el siguiente:

Administración de la primera fase de la escala NASA TLX. El experimento comenzaba con la realización de la primera fase de la escala NASA TLX en soporte de papel. Las instrucciones de dicha fase estaban escritas en la misma hoja de respuesta, aunque se les informaba de que si tenían alguna duda podían consultar al experimentador. En otra hoja estaban definidas las seis dimensiones de carga de trabajo de la escala NASA TLX (exigencia mental, exigencia física, exigencia temporal, esfuerzo, rendimiento y frustración). La tarea del sujeto consistió en valorar, por el método de comparación por pares de dimensiones, para una tarea indefinida, qué dimensión de cada par era más definitoria del concepto de carga de trabajo. Los resultados de esta primera fase se utilizan como ponderación de los resultados de la escala para cada tarea en la segunda fase.

Registro de las miradas. Se iniciaba con una fase de calibración del ojo derecho de cada su-

jeto para lo que se utilizaba una imagen proyectada con nueve puntos, coincidiendo con los márgenes de presentación de las letras en la pantalla.

Las condiciones experimentales. El experimento propiamente dicho comenzó con las instrucciones generales y la realización de las ocho condiciones experimentales. El sujeto controlaba a través de un teclado el comienzo de las instrucciones y de cada una de las condiciones experimentales. La mitad de los sujetos comenzaron con la condición Control (sólo tareas simples) y la otra mitad con la condición de Detectar (doble tarea), dentro de estas condiciones el orden de las tareas cognitivas fue contrabalanceado para cada sujeto. Después de escuchar las instrucciones de cada condición, comenzaba la presentación de letras en la pantalla y la realización simultánea de cada tarea. Cada condición tenía una duración de un minuto. En cada condición se proyectó como fondo visual una presentación simultánea de 16 letras mayúsculas que aparecían y desaparecían a una velocidad de una letra cada 125 ms., permaneciendo cada letra en pantalla durante 2 s. La presentación de las letras fue aleatorizada, tanto la letra que debía ser presentada como su posición en la pantalla.

Después de la realización de cada tarea y antes de escuchar las instrucciones de la condición siguiente, los sujetos completaban la segunda parte de la escala NASA TLX.

Se grabó de manera sincronizada, a 50 Hz/s, el diámetro de la pupila, las coordenadas de su mirada, el momento en que aparecía un estímulo y el momento en que respondía el sujeto. Para cada sujeto y condición se registraron 3000 líneas de datos.

RESULTADOS

Para cada variable dependiente se ha realizado un ANOVA de medidas repetidas (4 x 2) para la variable Tarea Cognitiva (con cuatro condiciones: Control (o no tarea), Escuchar, Produ-

cir palabras y Calcular) y la variable Demanda Visual (con dos condiciones: Sin detección y Detección; para todos los efectos en el texto nos podemos referir a estos dos niveles como tarea simple y doble tarea, respectivamente), con contrastes de cada tarea respecto a la condición control. Complementariamente se han realizado las comparaciones binarias entre las tareas, ajustando el error Tipo I por el método de Bonferroni. Se informa además del tamaño del efecto y de la potencia.

Para realizar los ANOVAs se comprobó el cumplimiento del supuesto de normalidad en cada distribución y se corrigieron los grados de libertad por Greenhouse- Geisser.

Además, se ha realizado un análisis de regresión sobre los aciertos en detección, usando como variables predictoras las tres medidas de carga, aún esperando que compartan varianza común.

NASA TLX

Los resultados de los juicios de los sujetos en la Escala NASA TLX a lo largo de las tareas pueden verse en la Figura 1.

Los resultados del ANOVA muestran diferencias estadísticamente significativas en las puntuaciones de los sujetos en la Escala NASA TLX para la variable Tarea Cognitiva, con $F(3, 141) = 121,25, p < 0,001$ ($\eta^2 = 0,721$ y potencia = 1). Los contrastes respecto a la condición Control muestran que existen diferencias significativas entre la condición de Control y Escuchar (30,00 vs. 47,52), entre la condición Control y Producir palabras (30,00 vs. 57,88) y entre la condición Control y Calcular (30,00 vs. 58,82). Las comparaciones por pares Bonferroni indican que todas las diferencias entre tareas son significativas $p < 0,001$ excepto la diferencia entre Calcular y Producir palabras, que son juzgadas iguales por los sujetos. La variable Demanda Visual muestra también diferencias estadísticamente significativas en los juicios de los sujetos, con mayor esfuerzo juzgado para la doble tarea que para la simple,

con $F(1,47) = 81,19, p < 0,001$ ($\eta^2 = 0,633$ y potencia = 1). La interacción entre Tarea y Demanda visual también es significativa y lo es para las tres tareas con respecto al Control. En los tres casos la interpretación es la misma: la diferencia por Demanda visual es menor cuando se realiza doble tarea que en las condiciones de control.

Pupila

Los resultados del tamaño de la pupila para cada tarea pueden verse en la Figura 2.

Los resultados del ANOVA muestran diferencias estadísticamente significativas en el Tamaño de la Pupila para la variable Tarea, con $F(3,141)=171,89, p < 0,001$ ($\eta^2 = 0,785$ y potencia = 1). Los contrastes respecto a la condición Control muestran que existen diferencias significativas entre la condición de Control y Escuchar (24,51 vs. 25,10), entre la Condición de Control y Producir palabras (24,51 vs. 27,18) y entre la Condición Control y Calcular (24,51 vs. 27,17). Según las comparaciones Bonferroni todas las diferencias entre medias son significativas excepto Producir palabras con Calcular. La variable Demanda Visual muestra diferencias estadísticamente significativas sobre el Tamaño de la Pupila, con $F(1,47)= 8,21, p = 0,006$ ($\eta^2 = 0,149$ y potencia = 0,801). La interacción Tarea por Demanda Visual es significativa, con $F(3,141)= 27,34, p < 0,001$ ($\eta^2 = 0,368$ y potencia = 1), en el sentido de que en las condiciones de poca carga (control y escuchar) la pupila se dilata más con detección que sin ella, mientras que en las condiciones con más carga (Producir palabras y Calcular) la pupila se dilata más sin detección, como si promediasen los efectos con los de la tarea de Detectar.

Parpadeo

Los resultados sobre el Parpadeo a lo largo de las tareas pueden verse en la Figura 3.

El ANOVA de medidas repetidas muestra diferencias estadísticamente significativas en el

Número de Parpadeos para la variable Tarea, con $F(3,141)=31,16$, $p<0,001$ ($\eta^2 = 0,399$ y potencia = 1). Los contrastes respecto a la condición Control muestran que existen diferencias significativas entre la condición Control y Escuchar (6,177 vs. 9,994), entre la condición Control y Producir palabras (6,177 vs. 15,969) y entre la condición Control y Calcular (6,177 vs. 16,469). El análisis Bonferroni indica que todas las diferencias entre medias son significativas, $p<0,01$, excepto al comparar Producir palabras con Calcular. En general, a mayor carga cognitiva más parpadeo. La variable Demanda Visual muestra también diferencias estadísticamente significativas sobre el Número de Parpadeos, con $F(1,47)=48,14$, $p<0,001$ ($\eta^2 = 0,506$ y potencia = 1), y en la dirección esperada: a mayor demanda visual (detección), menos parpadeo. La interacción entre Tarea y Demanda Visual no es significativa, por lo que podemos hablar de que el parpadeo es una medida sustractiva para esta combinación de tareas.

Detección de estímulos

Los resultados relativos a la proporción de aciertos en Detección, a lo largo de las tareas cognitivas pueden verse en la Figura 4.

El ANOVA de medidas repetidas muestra diferencias estadísticamente significativas en la proporción media de estímulos detectados en función de la Tarea, con $F(3,141)= 39,21$, $p<0,001$ ($\eta^2 = 0,445$ y potencia = 1). El contraste respecto a la condición de Control (Detectar) muestra un ligero descenso apenas significativo entre la condición de Detectar y

la condición de Detectar + Escuchar (0,810 vs. 0,701), un descenso mucho más acusado y significativo en la condición de Detectar + Producir palabras (0,810 vs. 0,425) y una diferencia de parecida magnitud entre la condición de Detectar y la condición de Detectar + Calcular (0,810 vs. 0,427). Las comparaciones por pares indican que la diferencia entre Producir palabras y Calcular no es estadísticamente significativa.

Regresión

Los coeficientes de regresión nos informan de la contribución de cada variable a la predicción del deterioro en la detección en la condición de doble tarea. En la Tabla 1 puede verse los resultados de la regresión (el método utilizado en la inclusión de las variables fue Introducir). En la condición de detección las tres medidas tienen contribuciones significativas. Para la condición sin detección la pupila es la variable que mayor varianza explica, y en la condición de detección la escala NASA TLX es la que mejor predice los aciertos. Probablemente la realización de doble tarea introduce factores no bien comprendidos que hacen que las medidas se comporten de forma diferente a como lo hacen cuando son tomadas con las tareas aisladas.

En la Tabla 2 se pueden ver las correlaciones entre las tres medidas de carga mental. La correlación más alta es entre NASA TLX y Pupila (0,717) cuando se realizan las tareas solas. Esta correlación desciende acusadamente al pasar a doble tarea.

Tabla 1. Coeficientes de regresión de las tres medidas de carga mental sobre los aciertos en Detección, según que las tareas se realicen sin o con Detección.

	R múltiple	Medida	Coef. estandarizado	t	Significación
Sin Detección	0,540	NASA TLX	-0,174	-1,971	0,05
		Pupila	-0,361	-4,070	0,000
		Parpadeo	-0,072	-1,011	0,313
Con Detección	0,627	NASA TLX	-0,291	-4,318	0,000
		Pupila	-0,238	-3,413	0,001
		Parpadeo	-0,259	-3,989	0,000

DISCUSIÓN Y CONCLUSIONES

En este experimento hemos tratado de analizar el comportamiento de tres medidas de carga mental, NASA TLX, Dilatación de la Pupila y Tasa de Parpadeo, tomadas durante la realización de tres tareas cognitivas (sin carga visual) y una de búsqueda visual continua. A su vez, esta tarea de búsqueda visual proporciona una cuarta medida objetiva de carga mental: el deterioro en la detección en situación de doble tarea. Las tareas cognitivas han sido, una de adquisición de información verbal (escuchar una descripción), otra de producción verbal y una tercera de producción continua de respuestas numéricas (también verbales) con restricciones aritméticas.

Considerando la carga de las tareas cognitivas, se puede concluir que: (a) todas ellas suponen carga mental respecto a la condición de control; (b) Producir palabras y Calcular son claramente iguales entre sí; (c) ambas producen una carga superior a Escuchar, que presenta resultados de carga mental débiles. Se replican los resultados de Recarte y Nunes (2003) en conducción real, con tareas similares y las mismas medidas de carga: la tarea de Escuchar implica baja carga cognitiva y la tarea de Producción verbal y Cálculo implica alta carga cognitiva y equivalente entre sí. Un apoyo a estos resultados lo hallamos también en el número de aciertos en Detección en la situación de doble tarea, hay una disminución de los aciertos relacionada con la carga mental de cada tarea. Así, esta disminución es apenas significativa para la tarea de Escuchar (con un tamaño del efecto de 0,09) y alta para la de Producción verbal y la de Cálculo (con un tamaño del efecto de 0,6).

Por lo que respecta a la tarea de búsqueda visual (comparada con el mero mirar a la misma matriz dinámica de estímulos) se produce una clara disociación: NASA TLX y Pupila indican que la detección sola produce una carga relativamente pequeña (algo superior a

Escuchar, pero inferior a Producir palabras y Calcular), pero el Parpadeo responde disminuyendo su tasa en la situación de detección, tal y como esperábamos en nuestra hipótesis, donde caracterizábamos esta disminución del parpadeo como un hecho inhibitorio, funcionalmente adecuado para no perder información en la tarea de búsqueda visual. NASA TLX y Pupila no discriminan cualitativamente entre carga cognitiva y carga visual, aunque el caso de la Pupila merece consideración aparte por la inesperada interacción que presenta la detección con las tareas cognitivas.

La pupila presenta una interacción entre la demanda visual (sin/con detección) y las tareas cognitivas que es contraintuitiva, porque para las tareas que implican más carga (Producir palabras y Calcular) se dilata más en la condición de tarea simple que en la de doble tarea. La mejor descripción de la interacción quizá sea ésta: la NASA TLX muestra la doble tarea como efectos aditivos, evaluando la carga de trabajo como la suma de la carga producida por cada componente (carga mental de la tarea cognitiva y carga visual de la detección). Es posible, incluso, que las personas no sean capaces de percibir la carga conjunta de las dos tareas y que el resultado sea una reconstrucción aditiva, una operación algorítmica, más que una percepción global. Por el contrario, la pupila, en el caso de tarea doble, muestra una tendencia a promediar las cargas de las dos tareas. Nuestra primera interpretación es que en la realización de estas dobles tareas, a diferencia de lo que ocurre cuando se realizan solas, la atención se reparte, probablemente como reparto de tiempo más que de cantidad de recursos atencionales, entre las dos tareas (Producir palabras y Detectar en un caso; Calcular y Detectar en el otro). Pero la tarea de Detectar implica menos carga mental que Producir palabras o que Calcular. Este reparto de la atención llevaría a momentos de menos carga mental (cuando la atención se presta a la detección) y otros de más carga mental (cuan-

do se presta a la tarea cognitiva). El resultado en la pupila sería un promedio de la carga a través del tiempo. Digamos que la detección supone una disminución o cierto descanso de procesamiento cognitivo. Efectos compatibles con esta hipótesis, aunque sin referencia a la pupila, sino con activación cerebral, pueden verse en Just et al. (2001).

En lo que respecta a la Tasa de Parpadeos, los resultados sostienen las hipótesis formuladas en la introducción. La tasa de parpadeos es capaz de distinguir entre demandas visuales y cognitivas de las tareas. Si comparamos la detección con sólo mirar a la pantalla, observamos un descenso significativo en el número de parpadeos cuando el sujeto está realizando tarea de detección (inhibición del parpadeo). Además, en las tareas simples, al comparar la condición de control (mirar la pantalla) con las tareas cognitivas, vemos un claro efecto de la tarea cognitiva en el sentido de que los parpadeos aumentan proporcionalmente a la carga mental que la tarea requiere (desinhibición del parpadeo). En las tareas dobles el número de parpadeos aumenta también respecto a la condición de control (Detección únicamente). Por consiguiente, podemos considerar confirmadas las hipótesis de Recarte y Nunes (2001) e insistir en la posibilidad de que el parpadeo se utilice como una medida peculiar capaz de distinguir entre la carga visual (inhibición) y la carga cognitiva (desinhibición). Desde el punto de vista aplicado, estos resultados son una base sólida para criticar el tratamiento del parpadeo, sin más matizaciones, como un indicador de fatiga, somnolencia o desactivación general, desarrollando dispositivos de aviso para los conductores.

Como recomendación aplicada hay que insistir en que se debería de tener sumo cuidado con la elección de la medida de carga mental en situaciones de doble tarea. Según nuestros resultados, existe una alta correlación entre las medidas utilizadas cuando las tareas que realizan los sujetos son simples. Sin

embargo, al combinar tareas puramente cognitivas (que no necesitan información visual) con una tarea de búsqueda visual continua, aparece una clara disociación y las tres medidas presentan patrones muy diferentes: La NASA TLX opera aditivamente con las dos tareas implicadas, la tasa de parpadeo refleja una sustracción y la pupila parece promediar los efectos de las dos tareas.

En suma, la disociación de las medidas en doble tarea parecen indicar que la Escala NASA TLX mide la carga de trabajo global, es decir, la cantidad de recursos totales invertidos (aunque dichos recursos sean de diferentes tipos), el Tamaño de la Pupila mide el reparto de atención, ya en tiempo o en recursos (i.e. nivel de activación) al realizar más de una tarea y la Tasa de Parpadeo mide, por un lado, la demanda de recursos visuales (a más demanda, menos parpadeo) y por otro, la cantidad de recursos cognitivos no visuales (a más carga cognitiva, más parpadeo).

REFERENCIAS

- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, 91, 276-292.
- De Waard, D. (1996). *The measurement of drivers' mental workload*. PhD thesis, University of Groningen. Haren, The Netherlands: University of Groningen, Traffic Research Centre.
- De Waard, D. (2002). Mental workload. In R. Fuller, & J.A. Santos (Eds.), *Human Factors for Highway Engineers* (pp. 161-176). Pergamon: Oxford.
- Fogarty, Ch., & Stern, J. A. (1989). Eye movements and blinks: their relationship to higher cognitive processes. *International Journal of Psychophysiology*, 8, 35-42.

- Hart, S.G., & Staveland, L.E. (1988). Development of NASA TLX (Task Load Index): results of empirical and theoretical research. In P.A. Hancock & N. Meshkati (Eds.), *Human Mental Workload* (pp. 139-183). Amsterdam: North-Holland.
- Hoecks, B., & Levelt, W. J. M. (1993). Pupillary dilatation as a measure of attention: A quantitative system analysis. *Behavior Research Methods, Instruments, & Computers*, 25, 16-26.
- Iqbal, S. T., Zheng, X. S., & Bailey, B. P. (2004). Task-Evoked Pupillary Response to Mental Workload in Human-Computer Interaction. In *CHI '04* (pp. 1477-1480). New York: AMC Press.
- Just, M.A., Carpenter, P., Keller, T., Emery, L., Zajac, H., & Thulborn, K. (2001). Interdependence of Nonoverlapping Cortical Systems in Dual Cognitive tasks. *NeuroImage*, 14, 417-426.
- Kubose, T., Bock, K., Dell, G.S., Garnsey, S. M., Kramer, A. F., & Mayhugh, J. (2006). The Effects of Speech Production and Speech Comprehension on Simulated Driving. *Performance Applied Cognitive Psychology*, 20(1), 43-63.
- Mulder, G. (1980). *The heart of mental effort*. PhD Thesis. Groningen: University of Groningen.
- Näätänen, R. (1992). *Attention and Brain Function*. Hillsdale, N. J.: LEA.
- O'Donnell, R.D., & Eggemeier, F.T. (1986). Workload assessment methodology. In K.R. Boff; L. Kaufman & J.P. Thomas (Eds.), *Handbook of perception and human performance*. Volumen II, cognitive processes and performance. New York: Wiley.
- Recarte, M. A., & Nunes, L. M. (2000). Effects of verbal and spatial imagery task on eye fixations while driving. *Journal of Experimental Psychology: Applied*, 6, 31-43.
- Recarte, M. A., & Nunes, L. M. (2001). Parpadeo durante la conducción: efectos de la carga mental y del tiempo conduciendo. *Vigilia-Sueño*, 14, 161-167.
- Recarte, M. A., & Nunes, L. M. (2003). Mental workload while driving: Effects on visual search, discrimination and decision making. *Journal of Experimental Psychology: Applied*, 9, 119-137.
- Rubio, S., Díaz, E., Martín, J., & Puente, J. M. (2004). Evaluation of subjective mental workload: a comparison of SWAT, NASA TLX and workload profile methods. *Applied Psychology*, 53, 61-86.
- Sanders, A. F. (1998). *Elements of Human Performance. Reaction Processes and attention in Human Skills*. Mahwah, NJ: Erlbaum.
- Stern, J. A., Boyer, D., & Schroeder, D. (1994). Blink rate: a possible measure of fatigue. *Human Factors*, 36, 285-297.
- Vicente, K.J., Thornton, D.C., & Moray, N. (1987). Spectral analysis of sinus arrhythmia: a measure of mental effort. *Human Factors*, 29, 171-182.
- Wickens, C.D. (1992). *Engineering psychology and human performance*. New York: HarperCollins.